



125

**PvuI (7)**  
**SgfI (6)** 1 GGATCTGGATCGCTCCGGTGCCCGTCAGTGGGAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTGGCAATTGAACGGGTGCCTA **MfeI (82)**  
101 GAGAAGTGGCGCGGGTAAACTGGAAAGTGATGTCGTGTAAGTGGTCCGCCTTTTCCGAGGGTGGGGGAGAACCCTATATAAGTGCAGTAGTCGCC

---

**Psp1406I (203)** 201 GTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGCTCGCATCTCTCCTTACGCGCCCGCCGCTACCTGAGGCC **HindIII (245)** **Bsu36I (291)**  
301 GCCATCCACGCGGGTTGAGTCGCGTTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAAGTGCCTCCGCGCTAGGTAAGTTAAAGCTCAGGTCGAGACC

---

401 GGGCCTTTGTCCGGCGCTCCCTTGAGCCTACCTAGACTCAGCCGGCTCTCCACGCTTTGCTGACCCTGCTTGTCAACTCTACGCTTTTGTTCGTTT **NgoMI (441)** **NaeI (441)**

---

501 TCTGTTCTGCGCCGTTACAGATCCAAGCTGTGACCGGGCGCTACCTGAGATCACCGGTACACCTGGGAGACGAGGATGAGGACGAGGGCTGCGCCGTGGA **NcoI (560)** **BstEII (555)** **AgeI (552)** **KasI (535)**  
1 M G D E D E D E G C A V E **XmnI (696)**

---

601 GCTGCAGATCACCGAAGCAACCTACCGGGCATGAGGAGAAGGTGAGCGTGGAGAATTCGCGCTGCTCAAGGTGCTGGGCACGGGAGCCTATGGGAAG **ScaI (953)**  
13> L Q I T E A N L T G H E E K V S V E N F A L L K V L G T G A Y G K  
701 GTGTTCTGGTGGGAAGACGGTGGGCACGACGCGGGCAAGCTCTATGCCATGAAGGTGCTACGCAAGGCGCGTGGTGCAGCGCGGAAGACACAGG  
47> V F L V R K T G G H D A G K L Y A M K V V L R K A A L V Q R A K T Q  
801 AGCATACCCGACCGAAGCTCGGTGCTGGAGCTGGTTCGKLAAGCACTTCTGTGACACTACGCTCCAGACGGATGCCAAGCTGCACCT  
80> E H T R T E R S V L E L V R Q A P F L V T L H Y A F Q T D A K L H L

---

901 CATCCTGGACTACGTGAGCGGTGGTGGATGTTCACTCACCTTACCAGCGCCAGTACTTCAAGGAGGCTGAGGTTGAGTGTATGGGGCGAGATTGTG  
113> I L D Y V S G G E M F T H L Y Q R Q Y F K E A E V R V Y G G E I V  
1001 CTGGCCCTGGAACACCTGCACAAGCTGGGTATCATCTACCGGGACCTGAAGCTGGAGAAGCTTACTTACTGACTCAGAAGTGCACATCGTCTTACAGACT  
147> L A L E H L H K L G I I Y R D L K L E N V L L D S E G H I V L T D  
1101 TTGGGCTGAGCAAGGAGTTCCTGACGGAGGAGAAGAGCGGACCTTCTCCTTCTGTGGACAATCGAGTACATGGCTCCCGAAATCATCCGAAGCAAGGC  
180> F G L S K E F L T E E K E R T F S F C G T I E Y M A P E I I R S K A  
1201 TGGACATGGCAAGGCTGTGGACTGGTGGAGCCTGGGTATCCTGCTCTTCGAGCTGCTGACAGGGGCTCACCTTACACTGGAGGGAGAGGAAACT  
213> G H G K A V D W W S L G I L L F E L L T G A S P F T L E G E R N T

---

1301 CAGGCTGAGGTGTCCGACGGATCTTGAAGTGCTCCCTCCCTCCCTCCCTCCGATTGGGCTGTGGCACAGGACCTGTACAGCGGCTGCTGTGCAAGG  
247> Q A E V S R R I L K C S P P F P L R I G P V A Q D L L Q R L L C K **BspEI (1349)**

---

1401 ACCCTAAGAAGAGTTGGGCGCAGGTCCCAGGGTGCAGGAAAGTCAAGAGTCAACCCCTTCTCCAGGGTCTGGACTGGGTTGGCTCTGGCTGCCAGAAA  
280> D P K K R L G A G P Q G A Q E V K S H P F F Q G L D W V A L A A R K **FspI (1434)**

---

1501 GATCCAGCCCCATTCCGGCCCCAGATCCGCTCAGAGCTGGATGTGGGAATTTTGGGAGGAATTCACCCGGCTGGAGCCGCTACTCCCTGCAGGC  
313> I P A P F R P Q I R S E L D V G N F A E E F T R L E P V Y S P A G **EcoRI (1561)** **SdaI (1591)**

---

1601 AGCCCTCCACCTGGGGACCTCGGATCTTTCAGGGATACTCCTCTGTTGCTCCGTCATCCTCTTTGACCACAACAATGCAGTGTGGCTGATGTACTGC  
347> S P P P G D P R I F Q G Y S F V A P S I L F D H N N A V M A D V L **SandI (1613)**

---

1701 AGGCACGGGTGCGGATACAGGCCCGGAGGGCAGCAGTGGCCAGGAGTGCATGATGCAGGACTCGCCTTTCTCCAGCAGTACGAACTGGACCTTCG  
380> Q A P G A G Y R P G R A A V A R S A M M Q D S P F F Q Q Y E L D L R **DraIII (1703)**

---

1801 GGAGCCAGCGTGGGGCAGGCTTCTCTGTGTGTCGGAGATGTAGGCAGCGCCAGAGCGGCCAGGAGTTTGTGTCAAGATCCTCAGCCGAGGCTG  
413> E P A L G Q G S F S V C R R C R Q R Q S G Q E F A V K I L S R R L **Eco47III (1806)**

---

1901 GAGGAGAACACTCAGCGAGAGGTGGCTGCTTTCGCCTGTGCCAGTACACCCCAACGTGGTGAATCTGCATGAGGTGCTTCATGACCAGCTACACACTT  
447> E E N T Q R E V A A L R L C Q S H P N V V N L H E V L H D Q L H T **BspHI (1980)**  
2001 ACCTGGTCTGGAGTTGCTGCGAGGCGGAGACTATTGGAACACATCCGCAAGAAGCGCTCTTCAGCGAGTCGGAGGCCAGCAGTCTTCGGAGCCT  
480> Y L V L E L L R G G E L L E H I R K K R L F S E S E A S Q I L R S L

---

2101 GGTTTCGGCGTGAGCTTATGCACGAGGAGGCGGCTGACCCGACCTGAAACCCGAGAACATCTTGTACGCGGACGACACTCCCGGGCCCCG  
513> V S A V S F M H E E A G V V H R D L K P E N I L Y A D D T P G A P **ApaLI (2140)** **XmaI (2188)**  
2201 GTGAAGATCATCGACTTCGGGTTTCGCGGACTGCGGCCAGAGCCCGGAGAGCCATGCAGACTCCTTGTTCACACTGCAGTACGCTGCACCCGAGC  
547> V K I I D F G F A R L R P Q S P A E P M Q T P C F T L Q Y A A P E

---

2301 TGCTGGCACAGCAGGGCTACGATGAGTCTGCGATCTATGGAGCCTGGGTGTCATTCTGTACATGATGCTGTCTGGCCAGGTTCCCTTCAAGGGCCTC  
580> L L A Q Q G Y D E S C D L W S L G V I L Y M M L S G Q V P F Q G A S **BsrGI (2357)**

---

2401 CGGCCAGGGTGGACAGAGTCAAGGAGTCAAGATCCGTGAAGGGCGCTTCTCCCTGGACGGGAAGCCTGGCAAGGTGTGTCGGAGGAA  
613> G Q G G Q S Q A A E I M C K I R E G R F S L D G E A W Q G V S E E **Tth111I (2411)**

---

2501 GCCAAGGAGCTGGTCCGAGGGCTACTGACAGTGCACCCGCCAAGCGGCTGAAGCTGGAGGGGCTGCGTAGCAGCTCGTGGCTTACAGGACGGCAGCGCGC  
647> A K E L V R G L L T V D P A K R L K L E G L R S S S W L Q D G S A **SalI (2530)**

2601 GCTCCTCGCCCCGCTCCGCACGCCGGATGTGCTGGAGTCTCTGGCCAGCTGTGCGTTCGGGCTCAATGCCACTTTCATGGCGTTCAACCGAGGCAA  
680▶ R S S P P L R T P D V L E S S G P A V R S G L N A T F M A F N R G K  
NotI (2773)  
2701 GCGCGAGGGCTTCTTTCTCAAGAGTGTAGAGAATGCGCCTCTGGCCAAGAGGCCAAGCAGAAGCTCCGGAGCGCGGCCCTCCGCGGAGGCTCCCCA  
713▶ R E G F F L K S V E N A P L A K R R K Q K L R S A A A S R R G S P  
NheI (2895)  
2801 GTGCTGCCTCCTCCGGTCGCCTACCAGCCTCTGCCGTAAGGGGACAACCTCGCCGAGCAACGGCCCTTGTCCCCTTCTAGCTCCCGCCACTGCTAG  
747▶ V P A S S G R L P A S A A K G T T R R A N G P L S P S •  
2901 CTGGCCAGACATGATAAGATACATTGATGAGTTTGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCT

---

3001 TTATTTGTAACCATTATAAGCTGCAATAAAACAGTTAAACAACAACAATTGCATTCAATTTATGTTTCAGGTTTCAGGGGGAGGTGTGGGAGGTTTTTTAA  
HpaI (3033) MfeI (3044)

---

3101 GCAAGTAAAACCTCTACAAATGTGGTATGGAATTCTAAAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGG  
EcoRI (3129)

---

3201 ATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTCTTTTCATGGAGTTTAAAGATATAGTGTATTTCCAA  
SspI (3368) SmaI (3382)

---

3301 GGTTTGAAGTACTCTTTCATTTCTTTATGTTTTAAATGCACTGACCTCCACATTCCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGC  
3401 AATGAAAAATAATGTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTT  
3501 TAATAGAAATTTGACAGCAAGAAAGCGAGCTTCTAGCTTTAGTTCCTGGTGTACTTGAGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCAT  
141▶ N R T Y K L P I L E E I T T K V L K G N  
SacI (3643)

---

3601 TCATCTCAATGAGCACAAAGCAGTCCAGGAGCATAGTCAGAGTCTGCACATGCCACAGGGGCTGACCACCCTGATGGACTGTCCACCTCATC  
120▶ M E I L V F C D P A Y D S I L E R C M G C P S V V R I S R D V E D  
3701 AGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAGTCTTCTGCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGTACCCTG  
87▶ S Y P H R V A V I T D F D K Q G N S V A S G I A I A E A C V T V R  
StuI (3807)  
Eco147I (3807)

---

3801 CCAATGTAGGCCTCAATGTGGACAGCAGAGATGATCTCCCGAGTCTTGGTCTGATGGCCGCCGACATGGTCTTGTGCTCATAGAGCATGGTGA  
53▶ G I Y A E I H V A S I I E G T K T R I A A G V H H K N D E Y L M T I  
BspHI (3957)

---

3901 TCTTCTCAGTGGCGACCTCCACCAGTCCAGATCCTGCTGAGAGATGTTGAAGGTCTTCTATGATGGCCCTCTATAGTGAAGTCTATTATACTATGCCGA  
20▶ K E T A V E V L E L D Q Q S I N F T K M  
XmnI (3949)

---

4001 TATACTATGCCGATGATTAATTGTCAAACAGCGTGGATGGCGTCTCCAGCTATCTGACGGTTCATAAACGAGCTCTGCTTATATAGACCTCCCACCG  
VspI (4015)  
AseI (4015) SacI (4072)

---

4101 TACACGCCTACCGCCATTTGCGTCAATGGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCCTTGATTTACTAGTCAAAACAACTCCCATTGACGTC  
SpeI (4170)

---

4200 AATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAACCGCTATCCAGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAAT  
Eco105I  
4300 ACGTAGATGACTGCCAAGTAGGAAAGTCCATAAGGTACTGTACTGGCATAATGCCAGGCGGGCCATTTACCGTATTGACGTCAATAGGGGGCGTAC  
NdeI (4403)

---

4400 TTGGCATATGATACACTTGTACTGCAAGTGGGCGAGTTTACCGTAAATACTCCACCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGG  
SdaI (4581) PacI (4589)

---

4500 AACATACGTCAATTTGACGTCAATGGGCGGGGTCGTTGGGCGTCCAGCCAGCGGGCCATTTACCGTAAAGTATGTAACGCC T G C A G G T T A A T T A A  
BspLU11I (4599)

---

4598 GAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCGCTTGTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAA  
4698 AATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCC  
4798 TGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCGTTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTCCG

---

4898 CTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCAACCCGGTAAGACACGACTTA  
ApaLI (4913)

---

4998 TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTA  
5098 GAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGG  
5198 TGGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTACGGGGTCTGACGCTCAGTGGAAACGAA

---

5298 AACTCACGTTAAGGGATTTTGGTATGGCTAGTTAATTAACATTTAAATC AGCGGCCGAATAAAATATCTTTATTTTTCATTACATCTGTGTGTTGGTTT  
PacI (5329) SmaI (5338) NotI (5348)

---

5398 TTTGTGTGAATCGTAACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAACTAGCAAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGA  
5498 ACATTTCTCTATCGAA